PRML PROJECT REPORT
NATURAL IMAGE CLASSIFICATION

Praneet Thakur (B19CSE066), Priyansh (B19CSE067), and Rahul Dhayal (B19EE069)

Abstract

There are many different technique and models to solve the problem of image classification.The purpose of this project is to gain a deeper understanding of different classification models, and how they perform on the CIFAR-10 dataset. As for the approach, we will implement various machine learning models(MLP,CNN and Rf) on the original dataset. For the second approach, we will perform dimensionality reduction using Principal Component Analysis (PCA) before applying the models.

Index Terms

Image Classification, CIFAR-10, PCA

I. Introduction

Image classification is one of the most fundamental problems in Machine Learning. It is the core foundation for bigger problems such as Computer Vision, Face Recognition System, or Self-driving car. With the development of deep Convolutional Neural Network (CNN), researchers have achieved good performance on the image recognition task. CIFAR-10 is a dataset consisting of 60,000 32x32 color images. We will compare the performance of CNN with classic machine learning models: MLP (Multilayer Perceptron), Random Forest. We will also explore the dimensionality reduction technique using PCA. The use of dimensionality reduction help reduce the training time of the models, while also can boost the accuracy. .

II. Dataset

Figure below shows the sample images from the CIFAR-10 dataset. This dataset consists of 60,000 32x32 color images [2]. There are also 10 labels in the dataset.

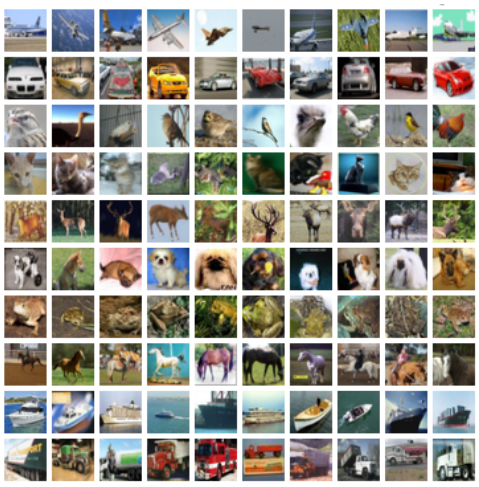

Fig. 1. Cifar-10 dataset sample images

III. System Design and Methodology

This section describes the models and methodology used in this project. The dataset is loaded from Keras' dataset. The data is already split into training and test set. Each pixel of the image is normalized to have value between 0 and 1..In A, the data is trained on the following models:

(1) MLP (Multilayer Perceptron) There are three hidden layers in the model one with 2050,1030,510 neurons each layer has a dropout layer of 0.2 in between.The optimizer used is adam and relu is used for activation.

(2) Random Forest- It uses sklearn library. The number of tress taken are 300.

(3) CNN-This function which uses keras library.Firstly there is a convolution network with two (4 cross 4) conv layers each layer has a maxpool (2*2) There are three hidden layers in the model one with 2050,1030,510 neurons each layer has a dropout

layer of 0.2 in between.The optimisise used is adam and relu is used for activation.
Module B uses PCA to extract features in the dataset before feeding into the models.The purpose of these approaches is to compare the performance of the models, and determine if dimensionality reduction techniques can help improve a model's performance. All the model configurations are same as the ones in without PCA.

A. Models without dimensionality reduction

Random Forest model is directly derived from Sklearn's library. MLP and CNN models are derived from the keras' library. For the CNN model, the performance is generally dependent on the depth of the network architecture. The deeper the network, the more likelihood that the CNN model perform better in image classification task. However, training a deep network requires high computing power and long training time.
The figure below shows the network architecture of the base line CNN model used for both dataset. The model uses Relu activation function and Adam optimizer. A dropout layer of 0.2 is applied after each deuse layer.

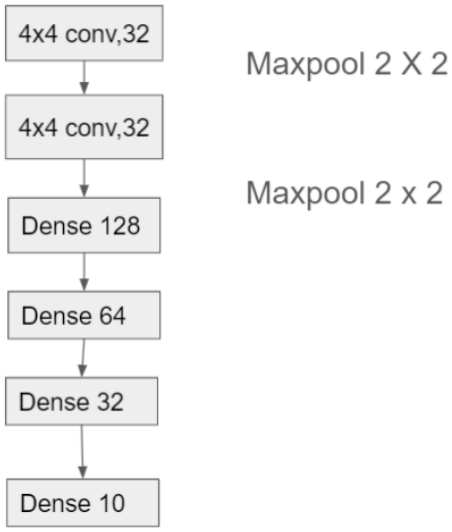

Fig. 2. Deep CNN model architecture

B. Dimensionality reduction with PCA

Dimensionality reduction technique used in this project is Principal Component Analysis (PCA).The features extracted is then trained on CNN, Random forest and Multilayer Perceptron (MLP).

PCA is a technique through which we can decompose the data into perpendicular vector where the spread of information of feature is more.As more variance means more information of data.

IV. Experiment and Analysis

This section summarises the results of the experiments and detailed analysis of each approach. To compare the performance of each model, the accuracy is used and cross validation scores is used.

A. Models without dimensionality reduction

For the baseline approach without using any dimensionality reduction, the training time is very long.For the CIFAR-10 dataset, the models need a lot of improvement.
All values in percentage.

Model	Accuracy	Cross-validation scores
MLP	43.58	42.26, 41.76, 45.2, 43.17, 44.02, 45.07
CNN	68.10	67.51, 67.55, 68.33, 68.73, 67.69, 68.76
Random forest	48.19	48.21, 48.18, 48.21 ,48.13, 47.94, 48.47

This Table summaries the experimental results for the baseline approach. The best model in this approach is CNN model, with a 68.10 percent accuracy for CIFAR-10 dataset.

B. dimensionality reduction with PCA

PCA is a dimensionality reduction technique which decreases the dimension of data by projecting it in the direction of maximum variance of the data. Using PCA, we can define upto what percentage of variance we want retained in the data.
All table values in percentage

Model	Accuracy	Cross-validation scores
MLP	55.50	55.20, 54.97, 56.45, 55.61, 55.76, 55.01
CNN	45.70	45.99, 45.30, 45.98 ,45.14, 46.03, 45.78
Random forest	44.07	43.46, 44.0, 43.84, 44.4, 44.19, 44.55

The Table 2 summarises the experimental results for the PCA dimensionality reduction approach. There is a slight decrease in performance for the Random Forest, and CNN model. The decrease in performance of the CNN model can be because of the result of PCA not preserving the original shape of the image.

It is also important to see that PCA reduces the performance of CNN with CIFAR-10 dataset a lot, from 68.10 percent down to 45.70 percent.This might be because PCA does not work well with color image with 3 channels. PCA dimensionality reduction also does not preserve the original image structure of the data, while CNN's input is an image.

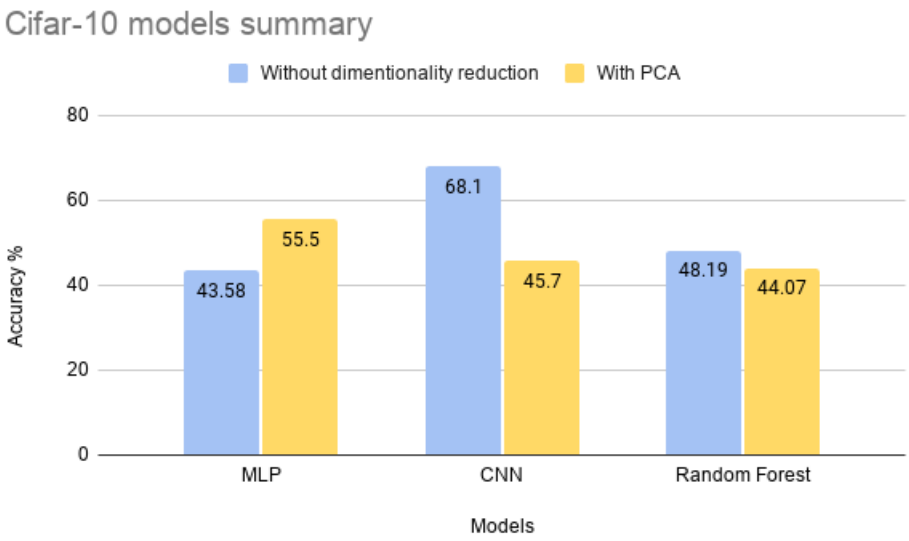

Fig. 3. Cifar-10 models summary

V. Conclusion

This report compares the performance of different models on the CIFAR-10 dataset. It also examines dimensionality reduction technique to boost the performance of the model.
On the basis of the results obtained, it can be concluded that the best model for cifar-10 dataset is CNN (without dimensionality reduction) MLP(with PCA).
The use of pretrained model can help improve the performance of all the models. Other dimensionality reduction and dimensionality reduction methods can also be used, such as using transfer learning,linear dimensionality reduction or CNN model to extract features.
The accuracy percentages obtained in all these models could have been higher by having a higher number of epochs depending on our computer system's configurations.

VI. Contribution

Research and planning- Praneet,Priyansh and Rahul
Coding- Priyansh and Praneet
Report- Praneet and Rahul
Readme file- Priyansh and Rahul

APPENDIX A

Label	Description
0	Airplane
1	Automobile
2	Bird
3	Cat
4	Deer
5	Dog
6	Frog
7	Horse
8	Ship
9	Truck

Fig. 4. Cifar-10 labels

REFERENCES

[1] https://keras.io/api/models/sequential/
https://keras.io/api/layers/convolution$_layers/convolution2d/$
$https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html$